## Why do we care about Distributions

Let us put it into the context of our final project

1. Consider the signboard with the text '**Mumbai**'. Now our classifier is analysing the text character by character, and a random variable <u>char</u> maps the character to one of the 26 possible characters in the english language

2. For the first character **M,** we know the True distribution intuitively.

| char | Y = P(char=c)<br>The certain event/True distribution | $\hat{Y}$<br>Obtained from model |
|------|------------------------------------------------------|-----------------------------------|
| a | 0 | 0.01 |
| b | 0 | 0.01 |
| ... | 0... | 0.01... |
| m | 1 | 0.7 |
| ... | ...0... | ...0.01... |
| z | ...0 | ...0.01 |

3. We compute the difference between the True and Predicted distributions using squared-error loss or some other loss function. From this, it is clear why we use distributions in the scope of our learning.